

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Korenberg, Julie R.
Yamakawa, Kazuhiro
- (ii) TITLE OF INVENTION: A NOVEL CHROMOSOME 21 GENE MARKER,
COMPOSITIONS AND METHODS USING SAME
- (iii) NUMBER OF SEQUENCES: 3
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: CAMPBELL & FLORES, LLP
 - (B) STREET: 4370 La Jolla Village Drive, Suite 700
 - (C) CITY: San Diego
 - (D) STATE: California
 - (E) COUNTRY: United States
 - (F) ZIP: 91212
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/337,690
 - (B) FILING DATE: 09-NOV-1994
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Campbell, Cathryn A.
 - (B) REGISTRATION NUMBER: 31,815
 - (C) REFERENCE/DOCKET NUMBER: P-CE 2573
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (619)535-9001
 - (B) TELEFAX: (619)535-8949

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5141 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: TRISOMY 21 FETAL BRAIN cDNA LIBRARY
 - (B) CLONE: EHOC-1
- (viii) POSITION IN GENOME:
 - (A) CHROMOSOME/SEGMENT: 21q22.3
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 157..3729

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTGCAGGAAT CGGCACGAGG CGGCGCAACC GGCTCCGGAG CTGCCTGGCG CGGCCGGGCG	60
GGCGGCGCCG CTCAGGCTCG GGCTCCGGCT GGGCCCGGCG CGGCCTCGGG GCTGCCCATG	120
GGGCGCGGGG GGCCGGGGCCG GTGACGCCGG ACGCCC ATG GAC GCC TCT GAG GAG	174
Met Asp Ala Ser Glu Glu	5
1	
CCG CTG CCG CCG GTG ATC TAC ACC ATG GAG AAC AAG CCC ATC GTC ACC	222
Pro Leu Pro Pro Val Ile Tyr Thr Met Glu Asn Lys Pro Ile Val Thr	20
10 15	
TGT GCT GGA GAT CAG AAT TTA TTT ACC TCT GTT TAT CCA ACG CTC TCT	270
Cys Ala Gly Asp Gln Asn Leu Phe Thr Ser Val Tyr Pro Thr Leu Ser	35
25 30	
CAG CAG CTT CCA AGA GAA CCA ATG GAA TGG AGA AGG TCC TAT GGC CGG	318
Gln Gln Leu Pro Arg Glu Pro Met Glu Trp Arg Arg Ser Tyr Gly Arg	50
40 45	
GCT CCG AAG ATG ATT CAC CTA GAG TCT AAC TTT GTT CAA TTC AAA GAG	366
Ala Pro Lys Met Ile His Leu Glu Ser Asn Phe Val Gln Phe Lys Glu	70
55 60	
GAG CTG CTG CCC AAA GAA GGA AAC AAA GCT CTG CTC ACG TTT CCC TTC	414
Glu Leu Leu Pro Lys Glu Gly Asn Lys Ala Leu Leu Thr Phe Pro Phe	85
75 80	
CTC CAT ATT TAC TGG ACA GAG TGC TGT GAT ACC GAA GTG TAT AAA GCT	462
Leu His Ile Tyr Trp Thr Glu Cys Cys Asp Thr Glu Val Tyr Lys Ala	100
90 95	
ACA GTA AAA GAT GAC CTC ACC AAG TGG CAG AAT GTT CTG AAG GCT CAT	510
Thr Val Lys Asp Asp Leu Thr Lys Trp Gln Asn Val Leu Lys Ala His	115
105 110	
AGC TCT GTG GAC TGG TTA ATA GTG ATA GTT GAA AAT GAT GCC AAG AAA	558
Ser Ser Val Asp Trp Leu Ile Val Ile Val Glu Asn Asp Ala Lys Lys	130
120 125	
AAA AAC AAA ACC AAC ATC CTT CCC CGA ACC TCT ATT GTG GAC AAA ATA	606
Lys Asn Lys Thr Asn Ile Leu Pro Arg Thr Ser Ile Val Asp Lys Ile	150
135 140	
AGA AAT GAT TTT TGT AAT AAA CAG AGT GAC AGG TGT GTT GTG CTC TCC	654
Arg Asn Asp Phe Cys Asn Lys Gln Ser Asp Arg Cys Val Val Leu Ser	165
155 160	
GAC CCC TTG AAG GAC TCT TCT CGA ACT CAG GAA TCC TGG AAT GCC TTC	702
Asp Pro Leu Lys Asp Ser Ser Arg Thr Gln Glu Ser Trp Asn Ala Phe	180
170 175	
CTG ACC AAA CTC AGG ACA TTG CTT CTT ATG TCT TTT ACC AAA AAC CTA	750
Leu Thr Lys Leu Arg Thr Leu Leu Leu Met Ser Phe Thr Lys Asn Leu	195
185 190	
GGC AAG TTT GAG GAT GAC ATG AGA ACC TTG AGG GAG AAG AGG ACT GAG	798
Gly Lys Phe Glu Asp Asp Met Arg Thr Leu Arg Glu Lys Arg Thr Glu	210
200 205	
CCA GGC TGG AGC TTT TGT GAA TAT TTC ATG GTT CAG GAG GAG CTT GCC	846
Pro Gly Trp Ser Phe Cys Glu Tyr Phe Met Val Gln Glu Glu Leu Ala	230
215 220	
225	

TTT	GTT	TTC	GAG	ATG	CTG	CAG	CAG	TTC	GAG	GAC	GCC	CTG	GTG	CAG	TAC	894
Phe	Val	Phe	Glu	Met	Leu	Gln	Gln	Phe	Glu	Asp	Ala	Leu	Val	Gln	Tyr	
				235					240					245		
GAC	GAA	CTG	GAC	GCC	CTC	TTC	TCT	CAG	TAT	GTG	GTC	AAC	TTC	GGG	GCC	942
Asp	Glu	Leu	Asp	Ala	Leu	Phe	Ser	Gln	Tyr	Val	Val	Asn	Phe	Gly	Ala	
			250					255					260			
GGG	GAT	GGT	GCC	AAC	TGG	CTG	ACT	TTT	TTC	TGC	CAG	CCA	GTG	AAG	AGC	990
Gly	Asp	Gly	Ala	Asn	Trp	Leu	Thr	Phe	Phe	Cys	Gln	Pro	Val	Lys	Ser	
		265					270					275				
TGG	AAC	GGA	TTG	ATC	CTC	CGA	AAA	CCC	ATA	GAT	ATG	GAG	AAG	CGG	GAA	1038
Trp	Asn	Gly	Leu	Ile	Leu	Arg	Lys	Pro	Ile	Asp	Met	Glu	Lys	Arg	Glu	
	280					285					290					
TCG	ATC	CAG	AGG	CGA	GAA	GCC	ACC	CTG	TTA	GAT	CTG	CGC	AGT	TAC	CTG	1086
Ser	Ile	Gln	Arg	Arg	Glu	Ala	Thr	Leu	Leu	Asp	Leu	Arg	Ser	Tyr	Leu	
					300					305					310	
TTC	TCT	CGC	CAG	TGC	ACC	TTG	CTG	CTC	TTC	CTG	CAG	AGG	CCG	TGG	GAG	1134
Phe	Ser	Arg	Gln	Cys	Thr	Leu	Leu	Leu	Phe	Leu	Gln	Arg	Pro	Trp	Glu	
				315					320					325		
GTG	GCC	CAG	CGC	GCC	CTA	GAG	CTG	CTG	CAC	AAC	TGC	GTG	CAG	GAA	CTG	1182
Val	Ala	Gln	Arg	Ala	Leu	Glu	Leu	Leu	His	Asn	Cys	Val	Gln	Glu	Leu	
			330				335						340			
AAG	CTC	TTA	GAA	GTC	TCT	GTC	CCA	CCT	GGT	GCT	CTG	GAC	TGC	TGG	GTG	1230
Lys	Leu	Leu	Glu	Val	Ser	Val	Pro	Pro	Gly	Ala	Leu	Asp	Cys	Trp	Val	
		345					350					355				
TTT	CTG	AGC	TGT	CTG	GAG	GTG	TTG	CAG	AGG	ATA	GAA	GGC	TGC	TGT	GAC	1278
Phe	Leu	Ser	Cys	Leu	Glu	Val	Leu	Gln	Arg	Ile	Glu	Gly	Cys	Cys	Asp	
	360					365					370					
CGG	GCA	CAG	ATC	GAC	TCA	AAC	ATT	GCC	CAC	ACT	GTG	GGG	CTA	TGG	AGC	1326
Arg	Ala	Gln	Ile	Asp	Ser	Asn	Ile	Ala	His	Thr	Val	Gly	Leu	Trp	Ser	
					375		380			385					390	
TAT	GCC	ACA	GAA	AAG	TTA	AAG	TCC	TTG	GGC	TAT	CTA	TGT	GGA	CTT	GTG	1374
Tyr	Ala	Thr	Glu	Lys	Leu	Lys	Ser	Leu	Gly	Tyr	Leu	Cys	Gly	Leu	Val	
				395					400					405		
TCA	GAG	AAA	GGA	CCT	AAC	TCA	GAA	GAT	CTC	AAC	AGG	ACA	GTT	GAC	CTT	1422
Ser	Glu	Lys	Gly	Pro	Asn	Ser	Glu	Asp	Leu	Asn	Arg	Thr	Val	Asp	Leu	
			410					415					420			
TTG	GCA	GGT	TTG	GGA	GCT	GAG	CGA	CCA	GAA	ACA	GCC	AAC	ACA	GCT	CAG	1470
Leu	Ala	Gly	Leu	Gly	Ala	Glu	Arg	Pro	Glu	Thr	Ala	Asn	Thr	Ala	Gln	
		425					430					435				
AGT	CCT	TAT	AAG	AAA	CTG	AAA	GAA	GCA	TTA	TCG	TCA	GTG	GAA	GCT	TTT	1518
Ser	Pro	Tyr	Lys	Lys	Leu	Lys	Glu	Ala	Leu	Ser	Ser	Val	Glu	Ala	Phe	
	440					445					450					
GAA	AAA	CAC	TAC	TTA	GAT	TTG	TCC	CAT	GCC	ACC	ATT	GAA	ATG	TAT	ACA	1566
Glu	Lys	His	Tyr	Leu	Asp	Leu	Ser	His	Ala	Thr	Ile	Glu	Met	Tyr	Thr	
	455				460					465					470	
AGC	ATT	GGG	AGG	ATT	CGA	TCT	GCT	AAG	TTT	GTT	GGA	AAA	GAT	CTG	GCA	1614
Ser	Ile	Gly	Arg	Ile	Arg	Ser	Ala	Lys	Phe	Val	Gly	Lys	Asp	Leu	Ala	
				475					480					485		
GAG	TTT	TAC	ATG	AGG	AAA	AAG	GCT	CCA	CAA	AAG	GCA	GAA	ATC	TAT	CTT	1662
Glu	Phe	Tyr	Met	Arg	Lys	Lys	Ala	Pro	Gln	Lys	Ala	Glu	Ile	Tyr	Leu	
			490					495					500			

CAA Gln	GGA Gly	GCA Ala 505	CTG Leu	AAA Lys	AAC Asn	TAC Tyr	CTG Leu 510	GCT Ala	GAG Glu	GGC Gly	TGG Trp	GCA Ala 515	CTC Leu	CCC Pro	ATC Ile	1710
ACA Thr	CAC His 520	ACA Thr	AGG Arg	AAG Lys	CAG Gln	CTG Leu 525	GCC Ala	GAA Glu	TGT Cys	CAA Gln 530	AAG Lys	CAC His	CTT Leu	GGA Gly	CAA Gln	1758
ATT Ile 535	GAA Glu	AAC Asn	TAC Tyr	CTG Leu	CAG Gln 540	ACC Thr	AGC Ser	AGC Ser	CTC Leu	TTA Leu 545	GCC Ala	AGT Ser	GAC Asp	CAC His	CAC His 550	1806
CTC Leu	ACT Thr	GAA Glu	GAG Glu	GAG Glu 555	CGC Arg	AAG Lys	CAC His	TTC Phe	TGC Cys 560	CAG Gln	GAG Glu	ATA Ile	CTT Leu	GAC Asp 565	TTT Phe	1854
GCC Ala	AGC Ser	CAG Gln 570	CCG Pro	TCA Ser	GAC Asp	AGC Ser	CCA Pro	GGT Gly 575	CAT His	AAG Lys	ATA Ile	GTG Val 580	CTA Leu	CCC Pro	ATG Met	1902
CAT His	TCC Ser	TTT Phe 585	GCA Ala	CAA Gln	CTG Leu	CGA Arg	GAT Asp 590	CTC Leu	CAT His	TTT Phe	GAT Asp 595	CCC Pro	TCC Ser	AAT Asn	GCC Ala	1950
GTG Val 600	GTC Val	CAC His	GTG Val	GGC Gly	GGC Gly	GTT Val 605	TTG Leu	TGC Cys	GTT Val	GAG Glu 610	ATA Ile	ACC Thr	ATG Met	TAC Tyr	AGC Ser	1998
CAG Gln 615	ATG Met	CCT Pro	GTG Val	CCT Pro	GTT Val 620	CAC His	GTG Val	GAG Glu	CAG Gln	ATT Ile 625	GTG Val	GTC Val	AAT Asn	GTC Val	CAC His 630	2046
TTC Phe	AGC Ser	ATT Ile	GAG Glu	AAA Lys 635	AAC Asn	AGC Ser	TAC Tyr	CGG Arg	AAG Lys 640	ACT Thr	GCG Ala	GAG Glu	TGG Trp	CTT Leu 645	ACC Thr	2094
AAG Lys	CAC His	AAG Lys 650	ACG Thr	TCC Ser	AAT Asn	GGG Gly	ATC Ile	ATT Ile 655	AAC Asn	TTT Phe	CCA Pro	CCC Pro	GAG Glu 660	ACC Thr	GCA Ala	2142
CCT Pro	TTC Phe	CCT Pro 665	GTA Val	TCC Ser	CAA Gln	AAC Asn	AGT Ser 670	TTG Leu	CCC Pro	GCG Ala	CTG Leu	GAG Glu 675	TTG Leu	TAT Tyr	GAA Glu	2190
ATG Met 680	TTT Phe	GAG Glu	AGA Arg	AGC Ser	CCA Pro	TCT Ser 685	GAT Asp	AAC Asn	TCC Ser	TTG Leu	AAC Asn 690	ACG Thr	ACT Thr	GGG Gly	ATT Ile	2238
ATC Ile 695	TGC Cys	AGA Arg	AAC Asn	GTC Val	CAC His 700	ATG Met	CTC Leu	CTG Leu	AGA Arg	AGG Arg 705	CAG Gln	GAG Glu	AGC Ser	AGC Ser	TCC Ser 710	2286
TCT Ser	CTA Leu	GAG Glu	ATG Met	CCC Pro 715	TCA Ser	GGG Gly	GTG Val	GCT Ala	CTG Leu 720	GAG Glu	GAG Glu	GGT Gly	GCC Ala	CAC His 725	GTG Val	2334
CTG Leu	AGG Arg	TGC Cys	AGC Ser	CAC His	GTG Val	ACC Thr	CTG Leu	GAA Glu 735	CCA Pro	GGG Gly	GCC Ala	AAC Asn	CAG Gln 740	ATA Ile	ACA Thr	2382
TTC Phe	AGG Arg	ACT Thr 745	CAG Gln	GCC Ala	AAG Lys	GAA Glu	CCT Pro 750	GGA Gly	ACG Thr	TAT Tyr	ACA Thr	CTC Leu 755	AGG Arg	CAG Gln	CTG Leu	2430
TGC Cys 760	GCC Ala	TCG Ser	GTG Val	GGC Gly	TCC Ser	GTG Val 765	TGG Trp	TTC Phe	GTC Val	CTC Leu	CCT Pro 770	CAC His	ATC Ile	TAC Tyr	CCC Pro	2478

ATT	GTG	CAG	TAC	GAC	GTG	TAC	TCA	CAG	GAG	CCC	CAG	CTG	CAC	GTG	GAG	2526
Ile	Val	Gln	Tyr	Asp	Val	Tyr	Ser	Gln	Glu	Pro	Gln	Leu	His	Val	Glu	
775					780					785					790	
CCG	CTG	GCT	GAT	AGC	CTT	CTG	GCA	GGC	ATT	CCT	CAG	AGA	GTC	AAG	TTC	2574
Pro	Leu	Ala	Asp	Ser	Leu	Leu	Ala	Gly	Ile	Pro	Gln	Arg	Val	Lys	Phe	
				795					800					805		
ACT	GTC	ACT	ACC	GGC	CAT	GAT	ACG	ATA	AAG	AAT	GGA	GAC	AGC	CTG	CAG	2622
Thr	Val	Thr	Thr	Gly	His	Asp	Thr	Ile	Lys	Asn	Gly	Asp	Ser	Leu	Gln	
			810					815					820			
CTT	AGC	AAT	GCC	GAA	GCC	ATG	CTC	ATC	CTG	TGC	CAG	GCG	GAG	AGC	AGG	2670
Leu	Ser	Asn	Ala	Glu	Ala	Met	Leu	Ile	Leu	Cys	Gln	Ala	Glu	Ser	Arg	
		825					830					835				
GCT	GTG	GTC	TAC	TCC	AAC	ACG	AGA	GAA	CAG	TCT	TCT	GAG	GCC	GCG	CTC	2718
Ala	Val	Val	Tyr	Ser	Asn	Thr	Arg	Glu	Gln	Ser	Ser	Glu	Ala	Ala	Leu	
	840					845					850					
CGG	ATT	CAG	TCC	TCC	GAC	AAG	GTC	ACG	AGC	ATC	AGT	CTG	CCT	GTT	GCG	2766
Arg	Ile	Gln	Ser	Ser	Asp	Lys	Val	Thr	Ser	Ile	Ser	Leu	Pro	Val	Ala	
855					860					865					870	
CCT	GCG	TAC	CAC	GTG	ATC	GAA	TTT	GAA	CTG	GAA	GTT	CTC	TCT	TTA	CCT	2814
Pro	Ala	Tyr	His	Val	Ile	Glu	Phe	Glu	Leu	Glu	Val	Leu	Ser	Leu	Pro	
				875					880					885		
TCA	GCC	CCA	GCA	CTC	GGA	GGG	GAG	AGT	GAC	ATG	CTG	GGG	ATG	GCA	GAG	2862
Ser	Ala	Pro	Ala	Leu	Gly	Gly	Glu	Ser	Asp	Met	Leu	Gly	Met	Ala	Glu	
			890					895					900			
CCC	CAC	AGG	AAG	CAT	AAG	GAC	AAA	CAG	AGA	ACT	GGC	CGC	TGC	ATG	GTT	2910
Pro	His	Arg	Lys	His	Lys	Asp	Lys	Gln	Arg	Thr	Gly	Arg	Cys	Met	Val	
		905					910					915				
ACC	ACA	GAC	CAC	AAA	GTG	TCG	ATT	GAC	TGC	CCG	TGG	TCC	ATC	TAC	TCC	2958
Thr	Thr	Asp	His	Lys	Val	Ser	Ile	Asp	Cys	Pro	Trp	Ser	Ile	Tyr	Ser	
	920					925					930					
ACA	GTC	ATC	GCA	CTG	ACC	TTC	AGC	GTA	CCC	TTC	AGG	ACC	ACA	CAC	AGC	3006
Thr	Val	Ile	Ala	Leu	Thr	Phe	Ser	Val	Pro	Phe	Arg	Thr	Thr	His	Ser	
935					940					945					950	
CTC	CTG	TCC	TCA	GGA	ACA	CGG	AAA	TAT	GTT	CAA	GTT	TGT	GTC	CAG	AAT	3054
Leu	Leu	Ser	Ser	Gly	Thr	Arg	Lys	Tyr	Val	Gln	Val	Cys	Val	Gln	Asn	
				955					960					965		
TTG	TCA	GAA	CTT	GAC	TTT	CAG	CTG	TCA	GAT	AGT	TAT	CTT	GTA	GAT	ACC	3102
Leu	Ser	Glu	Leu	Asp	Phe	Gln	Leu	Ser	Asp	Ser	Tyr	Leu	Val	Asp	Thr	
			970					975					980			
GGT	GAT	AGT	ACC	GAC	CTG	CAA	CTA	GTA	CCA	CTG	AAC	ACG	CAG	TCC	CAG	3150
Gly	Asp	Ser	Thr	Asp	Leu	Gln	Leu	Val	Pro	Leu	Asn	Thr	Gln	Ser	Gln	
		985					990					995				
CAG	CCC	ATC	TAC	AGC	AAG	CAG	TCG	GTG	TTC	TTC	GTC	TGG	GAA	CTC	AAG	3198
Gln	Pro	Ile	Tyr	Ser	Lys	Gln	Ser	Val	Phe	Phe	Val	Trp	Glu	Leu	Lys	
	1000					1005					1010					
TGG	ACA	GAA	GAG	CCT	CCC	CCT	TCT	CTG	CAT	TGC	CGG	TTC	TCT	GTT	GGA	3246
Trp	Thr	Glu	Glu	Pro	Pro	Pro	Ser	Leu	His	Cys	Arg	Phe	Ser	Val	Gly	
1015					1020					1025					1030	
TTT	TCC	CCA	GCT	TCT	GAG	GAA	CAG	CTG	TCT	ATC	TCC	TTA	AAG	CCG	TAT	3294
Phe	Ser	Pro	Ala	Ser	Glu	Glu	Gln	Leu	Ser	Ile	Ser	Leu	Lys	Pro	Tyr	
				1035					1040					1045		

ACT TAT GAA TTT AAA GTG GAA AAT TTT TTT ACA TTA TAC AAC GTG AAG	3342
Thr Tyr Glu Phe Lys Val Glu Asn Phe Phe Thr Leu Tyr Asn Val Lys	
1050 1055 1060	
GCT GAG ATC TTT CCC CCT TCG GGA ATG GAG TAT TGC AGA ACA GGC TCC	3390
Ala Glu Ile Phe Pro Pro Ser Gly Met Glu Tyr Cys Arg Thr Gly Ser	
1065 1070 1075	
CTC TGC TCC CTG GAG GTT TTG ATC ACG AGG CTC TCA GAC CTC TTG GAG	3438
Leu Cys Ser Leu Glu Val Leu Ile Thr Arg Leu Ser Asp Leu Leu Glu	
1080 1085 1090	
GTG GAT AAA GAT GAA GCA CTG ACT GAA TCT GAT GAG CAT TTT TCG ACA	3486
Val Asp Lys Asp Glu Ala Leu Thr Glu Ser Asp Glu His Phe Ser Thr	
1095 1100 1105 1110	
AAG CTT ATG TAT GAA GTT GTC GAC AAC AGT AGC AAC TGG GCA GTG TGT	3534
Lys Leu Met Tyr Glu Val Val Asp Asn Ser Ser Asn Trp Ala Val Cys	
1115 1120 1125	
GGG AAA AGC TGC GGT GTC ATC TCC ATG CCA GTG GCT GCT CGG GCC ACT	3582
Gly Lys Ser Cys Gly Val Ile Ser Met Pro Val Ala Ala Arg Ala Thr	
1130 1135 1140	
CAC AGG GTC CAC ATG GAA GTG ATG CCG CTC TTC GCC GGG TAT CTC CCC	3630
His Arg Val His Met Glu Val Met Pro Leu Phe Ala Gly Tyr Leu Pro	
1145 1150 1155	
CTG CCC GAC GTC AGG CTG TTC AAG TAC CTC CCC CAT CAT TCT GCA CAC	3678
Leu Pro Asp Val Arg Leu Phe Lys Tyr Leu Pro His His Ser Ala His	
1160 1165 1170	
TCC TCC CAA CTG GAC GCT GAC AGC TGG ATA GAA AAC GCA GCC TGT CAG	3726
Ser Ser Gln Leu Asp Ala Asp Ser Trp Ile Glu Asn Ala Ala Cys Gln	
1175 1180 1185 1190	
TAGACAAGCA CGGGGACGAC CAGCCGGACA GCAGCAGCCT CAAGAGCAGG GGCAGCGTGC	3786
ATTCGGCCTG CAGCAGCGAG CACAAAGGCC TACCCATGCC CCGGCTGCAG GCACTGCCGG	3846
CCGGCCAGGT CTTCAACTCC AGCTCGGGCA CACAAGTCCT GGTCATCCCC AGCCAAGATG	3906
ACCACGTCCT GGAAGTCAGT GTAACATGAC AACGCCAGGG TGAACACACG CCACTTCCCA	3966
GCTAGGAGTG CACTTTATGG GACTGTGACT GGACTCTTCC GTTCTGGCTC CAGCCAGACC	4026
TTCAGTGGTC CTGCCTGGCC GTGGGGACAT CAGAGAGTGT CATCACGCAG CTGGCCAGCT	4086
GAGTTCTGTT GTTGTTTTCA TGCCGCCTGT GATCTCAGAT TCCTGCTTTT CTCACCCCGT	4146
CCCCATGCTG GTGTCCGACG CCGCTTACTC AGAGCCCTGG CCTCCCTCCC CCTACCTCAC	4206
ACGCTGCTCA TGAAAGTTTC CACCCACGCT GTCTCCACGG AACAGCCTCC GTCTGCTGGC	4266
TCTTCGTGGA AGGCCATTTG TCTTTCAGGT AGACACTCAG CAGCCCTCAC GGTCTTAGTG	4326
ACGTGTGTGC CTTTCTGGTC ACACAGCTGC CCAGTTTCCT GATCGGGGTG GATTTGTGTC	4386
CCCTAAGGGG TAAACAGCC GTTTACCGCA GATCCTCTCA TACACCCTTC TAGGGGAGGC	4446
GGGTGGGGGA GGGAGGGATC ATAACCCCTT CTGTGCCTTG GGATGCCGGA GCTGGGGGAC	4506
CTGGAGGCCC ATCAGCCGGA GCCACGTGAA AGGTACTGAA GAAAGCTGAG ACCCGGCTGT	4566
GAGGAGCGCC TCAGCGGTGA GGTGGTTTAG GGATAAATGT TTCTGGAACC CTGTGGTCCC	4626
CCATAATGTT GATAGATATC ATATGCACTG GGAGTTAAAT ATATTTAATT TAATGATCAT	4686

TATATATGTG GGGGTTAATA TGTTGTTTTT CTGTCCCTTT AAAGTCTTTA CATGTAATTG 4746
 TAGCTGTATA ATCGTTATTT TTCTTTTGCA TCTTAAGTCT TAGAAATTAA GATATTCCAT 4806
 CGTGAGGATG AGAGAGGTCC TCA GTGTGTT TTTGGTCTGG TTGTAGGGAA GGA CTCAAGT 4866
 CCTGGAATGT CCTCCACTGG TCTACTGAGT TGCAGTCACA CTGTTCCAAT GGATTATTTG 4926
 CTTTCGGTTG TAAATTTAAT TGTACATATG GTTGATTAT TATTTTAA AATACAGACT 4986
 AACTGATGTA ATGTTTATGT ATAAGTTGCA CCAAAAATCA AGGACAAAAA TAAGTGTGTT 5046
 TGTTTTTACA GGTGTGAAAG TCACAGCTTG TAAATAAGTG TTGTATGTAT TAAACCTTTT 5106
 CCAGTCTCTCC AAAAAAAAAA AAAAAAAAAA AAAAA 5141

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1190 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Ala Ser Glu Glu Pro Leu Pro Pro Val Ile Tyr Thr Met Glu
 1 5 10 15
 Asn Lys Pro Ile Val Thr Cys Ala Gly Asp Gln Asn Leu Phe Thr Ser
 20 25 30
 Val Tyr Pro Thr Leu Ser Gln Gln Leu Pro Arg Glu Pro Met Glu Trp
 35 40 45
 Arg Arg Ser Tyr Gly Arg Ala Pro Lys Met Ile His Leu Glu Ser Asn
 50 55 60
 Phe Val Gln Phe Lys Glu Glu Leu Leu Pro Lys Glu Gly Asn Lys Ala
 65 70 75 80
 Leu Leu Thr Phe Pro Phe Leu His Ile Tyr Trp Thr Glu Cys Cys Asp
 85 90 95
 Thr Glu Val Tyr Lys Ala Thr Val Lys Asp Asp Leu Thr Lys Trp Gln
 100 105 110
 Asn Val Leu Lys Ala His Ser Ser Val Asp Trp Leu Ile Val Ile Val
 115 120 125
 Glu Asn Asp Ala Lys Lys Lys Asn Lys Thr Asn Ile Leu Pro Arg Thr
 130 135 140
 Ser Ile Val Asp Lys Ile Arg Asn Asp Phe Cys Asn Lys Gln Ser Asp
 145 150 155 160
 Arg Cys Val Val Leu Ser Asp Pro Leu Lys Asp Ser Ser Arg Thr Gln
 165 170 175
 Glu Ser Trp Asn Ala Phe Leu Thr Lys Leu Arg Thr Leu Leu Leu Met
 180 185 190
 Ser Phe Thr Lys Asn Leu Gly Lys Phe Glu Asp Asp Met Arg Thr Leu
 195 200 205

003327 00000000

Arg	Glu	Lys	Arg	Thr	Glu	Pro	Gly	Trp	Ser	Phe	Cys	Glu	Tyr	Phe	Met
210						215					220				
Val	Gln	Glu	Glu	Leu	Ala	Phe	Val	Phe	Glu	Met	Leu	Gln	Gln	Phe	Glu
225					230					235					240
Asp	Ala	Leu	Val	Gln	Tyr	Asp	Glu	Leu	Asp	Ala	Leu	Phe	Ser	Gln	Tyr
				245					250					255	
Val	Val	Asn	Phe	Gly	Ala	Gly	Asp	Gly	Ala	Asn	Trp	Leu	Thr	Phe	Phe
			260					265					270		
Cys	Gln	Pro	Val	Lys	Ser	Trp	Asn	Gly	Leu	Ile	Leu	Arg	Lys	Pro	Ile
		275					280					285			
Asp	Met	Glu	Lys	Arg	Glu	Ser	Ile	Gln	Arg	Arg	Glu	Ala	Thr	Leu	Leu
	290					295					300				
Asp	Leu	Arg	Ser	Tyr	Leu	Phe	Ser	Arg	Gln	Cys	Thr	Leu	Leu	Leu	Phe
305					310					315					320
Leu	Gln	Arg	Pro	Trp	Glu	Val	Ala	Gln	Arg	Ala	Leu	Glu	Leu	Leu	His
				325					330					335	
Asn	Cys	Val	Gln	Glu	Leu	Lys	Leu	Leu	Glu	Val	Ser	Val	Pro	Pro	Gly
			340					345					350		
Ala	Leu	Asp	Cys	Trp	Val	Phe	Leu	Ser	Cys	Leu	Glu	Val	Leu	Gln	Arg
		355					360					365			
Ile	Glu	Gly	Cys	Cys	Asp	Arg	Ala	Gln	Ile	Asp	Ser	Asn	Ile	Ala	His
	370					375					380				
Thr	Val	Gly	Leu	Trp	Ser	Tyr	Ala	Thr	Glu	Lys	Leu	Lys	Ser	Leu	Gly
385					390					395					400
Tyr	Leu	Cys	Gly	Leu	Val	Ser	Glu	Lys	Gly	Pro	Asn	Ser	Glu	Asp	Leu
				405					410					415	
Asn	Arg	Thr	Val	Asp	Leu	Leu	Ala	Gly	Leu	Gly	Ala	Glu	Arg	Pro	Glu
			420					425					430		
Thr	Ala	Asn	Thr	Ala	Gln	Ser	Pro	Tyr	Lys	Lys	Leu	Lys	Glu	Ala	Leu
		435					440					445			
Ser	Ser	Val	Glu	Ala	Phe	Glu	Lys	His	Tyr	Leu	Asp	Leu	Ser	His	Ala
	450					455					460				
Thr	Ile	Glu	Met	Tyr	Thr	Ser	Ile	Gly	Arg	Ile	Arg	Ser	Ala	Lys	Phe
465					470					475					480
Val	Gly	Lys	Asp	Leu	Ala	Glu	Phe	Tyr	Met	Arg	Lys	Lys	Ala	Pro	Gln
				485					490					495	
Lys	Ala	Glu	Ile	Tyr	Leu	Gln	Gly	Ala	Leu	Lys	Asn	Tyr	Leu	Ala	Glu
			500					505					510		
Gly	Trp	Ala	Leu	Pro	Ile	Thr	His	Thr	Arg	Lys	Gln	Leu	Ala	Glu	Cys
		515					520					525			
Gln	Lys	His	Leu	Gly	Gln	Ile	Glu	Asn	Tyr	Leu	Gln	Thr	Ser	Ser	Leu
	530					535					540				
Leu	Ala	Ser	Asp	His	His	Leu	Thr	Glu	Glu	Glu	Arg	Lys	His	Phe	Cys
545					550					555					560

Gln	Glu	Ile	Leu	Asp 565	Phe	Ala	Ser	Gln	Pro 570	Ser	Asp	Ser	Pro	Gly 575	His
Lys	Ile	Val	Leu 580	Pro	Met	His	Ser	Phe 585	Ala	Gln	Leu	Arg	Asp 590	Leu	His
Phe	Asp	Pro 595	Ser	Asn	Ala	Val	Val 600	His	Val	Gly	Gly	Val 605	Leu	Cys	Val
Glu	Ile 610	Thr	Met	Tyr	Ser	Gln 615	Met	Pro	Val	Pro	Val 620	His	Val	Glu	Gln
Ile 625	Val	Val	Asn	Val	His 630	Phe	Ser	Ile	Glu	Lys 635	Asn	Ser	Tyr	Arg	Lys 640
Thr	Ala	Glu	Trp 645	Leu	Thr	Lys	His	Lys	Thr 650	Ser	Asn	Gly	Ile	Ile 655	Asn
Phe	Pro	Pro	Glu 660	Thr	Ala	Pro	Phe	Pro 665	Val	Ser	Gln	Asn	Ser 670	Leu	Pro
Ala	Leu	Glu 675	Leu	Tyr	Glu	Met	Phe 680	Glu	Arg	Ser	Pro	Ser 685	Asp	Asn	Ser
Leu	Asn 690	Thr	Thr	Gly	Ile	Ile 695	Cys	Arg	Asn	Val	His 700	Met	Leu	Leu	Arg
Arg 705	Gln	Glu	Ser	Ser	Ser 710	Ser	Leu	Glu	Met	Pro 715	Ser	Gly	Val	Ala	Leu 720
Glu	Glu	Gly	Ala	His 725	Val	Leu	Arg	Cys	Ser 730	His	Val	Thr	Leu	Glu 735	Pro
Gly	Ala	Asn	Gln 740	Ile	Thr	Phe	Arg	Thr 745	Gln	Ala	Lys	Glu	Pro 750	Gly	Thr
Tyr	Thr	Leu 755	Arg	Gln	Leu	Cys	Ala 760	Ser	Val	Gly	Ser	Val 765	Trp	Phe	Val
Leu	Pro 770	His	Ile	Tyr	Pro	Ile 775	Val	Gln	Tyr	Asp	Val 780	Tyr	Ser	Gln	Glu
Pro 785	Gln	Leu	His	Val	Glu 790	Pro	Leu	Ala	Asp	Ser 795	Leu	Leu	Ala	Gly	Ile 800
Pro	Gln	Arg	Val	Lys 805	Phe	Thr	Val	Thr	Thr 810	Gly	His	Asp	Thr	Ile 815	Lys
Asn	Gly	Asp	Ser 820	Leu	Gln	Leu	Ser	Asn 825	Ala	Glu	Ala	Met	Leu 830	Ile	Leu
Cys	Gln	Ala 835	Glu	Ser	Arg	Ala	Val 840	Val	Tyr	Ser	Asn	Thr 845	Arg	Glu	Gln
Ser	Ser 850	Glu	Ala	Ala	Leu	Arg 855	Ile	Gln	Ser	Ser	Asp 860	Lys	Val	Thr	Ser
Ile 865	Ser	Leu	Pro	Val	Ala 870	Pro	Ala	Tyr	His	Val 875	Ile	Glu	Phe	Glu	Leu 880
Glu	Val	Leu	Ser	Leu 885	Pro	Ser	Ala	Pro	Ala 890	Leu	Gly	Gly	Glu	Ser 895	Asp
Met	Leu	Gly	Met 900	Ala	Glu	Pro	His	Arg 905	Lys	His	Lys	Asp 910	Lys	Gln	Arg

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Moraxella bovis
- (C) INDIVIDUAL ISOLATE: MboI linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCTGATGCTC GAGTGAATTC